

# Diseño de un sistema de información y evaluación científica (Doctoral Thesis by Daniel Torres-Salinas)

Por Henk F. Moed

**Moed, Henk F.** "Diseño de un sistema de información y evaluación científica (Doctoral Thesis by Daniel Torres Salinas)". En: *El profesional de la información*, 2007, noviembre-diciembre, v. 16, n. 6, pp. 661-663.

DOI: 10.3145/epi.2007.nov.14



Henk F. Moed y Daniel Torres Salinas

**QUANTITATIVE STUDY OF SCIENCE AND TECHNOLOGY is a rapidly developing field, and its development is closely linked to a number of general tendencies in scholarship throughout the world. National governments and research organisations and institutions need systematic evaluations to optimise their research allocations, re-orient their research support, justify research organisations, restructure research in particular fields, or augment research productivity.**

Evaluative bibliometrics is a subfield of quantitative science and technology studies, intended to construct indicators of research performance from a quantitative analysis of scholarly documents. Citation analysis, one of its key methodologies, assesses the contributions scholars make in their research publications to the advancement of scholarly knowledge. During the past decades, numerous studies applied and further developed

citation analysis in the assessment of scientific research performance. They reached the conclusion that bibliometric indicators can assist in building up insight into the quality of scholarly work under evaluation and in forming a judgment, and hence constitute a research evaluation *tool*, provided that they have a high level of sophistication and are derived from accurate bibliometric data.

This notion plays a key role in the dissertation of **Daniel Torres-Salinas**, entitled *Design of a system of scientific information and evaluation. A scientometric analysis of the research activities at the University of Navarra in the health sciences (1999-2005)*; the Spanish title is: *Diseño de un sistema de información y evaluación científica. Análisis cuantitativo de la actividad investigadora de la Universidad de Navarra en el área de ciencias de la salud. (1999-2005)*. Supervisors were **Evaristo Jiménez-Contreras** and **Emilio Delgado-López-Có-**

**zar** at the *University of Granada*. It presents the development of an information system on research activities in a scientific-scholarly institution that aims at providing a useful tool in research evaluation and management. The author carried out his work along the following five lines:

- The theoretical specification of such an information system and the type of science indicators included therein, in close interaction with potential users of the system.

- Its technical realization, including the development of various types of data entry processes, -partly manual, partly automated-.

- The actual application of the system to a particular case: the *University of Navarra*, a private Spanish university specializing in health sciences.

- An analysis of the results in terms of their implications for research performance at the *University of Navarra*.

- A discussion of his empirical findings from the perspective of more general research issues in the field of quantitative science studies and research assessment.

Life sciences departments at the *University of Navarra* constituted the object of the study. Data were collected from this university's internal databases, especially those on grant-funded projects and on scientific personnel, and from two major databases of scientific literature: the *Web of Science* published

by *Thomson Scientific* (formerly the *Institute for Scientific Information*) and *Scopus*, compiled by *Elsevier*.

The system includes both input and output indicators. The principal aspects of the output of scientific work covered by the system are: production, based on the number of published articles; impact or visibility at the international research front, derived from citation counts; and scientific collaboration, as reflected in co-authorship in scientific publications. Main aspects at the input side are the research capacity, measured by the number of active researchers, and the amount of external funding. Indicators were calculated at various levels of aggregation: for the institution as a whole and per department, thematic subfield and individual researcher.

The dissertation presents numerous interesting methodologies, results, observations and conclusions. The data tables and especially the figures are of very high quality. Below I highlight the study's outcomes and conclusions of a more general nature that relate to three issues, currently receiving much attention in the field of quantitative science studies. The first deals with scientific literature databases: how do indicators derived from *Thomson Scientific's Web of Science (WoS)* compare to those obtained from *Elsevier's Scopus*? A second issue is: does scientific collaboration pay? More specifically: how does the citation impact of research articles resulting from collaboration—especially international collaboration—compare to that of non-collaborative papers? Finally, I address the analysis of the relationship between 'input' and 'output' of scientific research activity.

### 1. Web of Science versus Scopus

One of the objectives of Daniel Torres Salinas' study was to obtain insight into differences between

citation impact indicators based on *WoS* data and those derived from *Scopus*. *Scopus* is a new multi-disciplinary citation index published by Elsevier, covering some 15.000 sources, mostly in science, technology and medicine. Although a number of recent studies compared *Scopus* and *WoS*, only a very few compared citation counts derived from the two databases. Torres Salinas' dissertation presents such a comparison. His comparative analysis of *Scopus* and *WoS* included citation counts for about 2.300 articles published from the University of Navarra and included in the *WoS*. Papers in *Scopus* journals not covered by the *WoS* were not taken into account. It was found that *Scopus* has almost 15 per cent more citations to these papers than the *WoS*.

Focusing on the largest subfields in terms of number of published articles, the comparative percentage of citations was somewhat higher in neurosciences, neurology and neurosurgery, hematology and hematotherapy, preventive medicine and public health, and endocrinology and nutrition, and below the average (but nevertheless above zero) in internal medicine, gen therapy, oncology, biochemistry and molecular biology, genetics, and microbiology and parasitology. The ranking of departments based on *Scopus* citation counts was similar to that based on *WoS* citation data. Out of 50 departments, 20 were at the same position, 17 were higher in the *Scopus* ranking (13 by one position only), and 13 were lower (7 by one position).

**Torres-Salinas** makes the significant observation that, even though *Scopus* covers about 6.000 (or 65 per cent) more journals than the *WoS*, the citation counts to the papers published from the *University of Navarra* derived from *Scopus* are only some 15 per cent higher than those based on the *WoS*. He suggests that this discrepan-

cy is due to differences in citation circuits between core and peripheral journals, assuming that *Scopus* contains more peripheral journals than the *WoS*.

As a rule, the use of multiple databases provides a more complete picture; more insight is needed into differences in coverage between *Scopus* and *WoS* and in citation patterns between core and peripheral journals. The outcomes of the study presented in **Torres Salinas'** dissertation constitute a proper starting point for further analysis, and point towards the possibility to carry out future bibliometric studies of citation impact based upon *Scopus* data only.

### 2. Does scientific collaboration pay?

The statistical relationship between scientific collaboration and citation impact, and the effects of collaboration on citation impact and vice versa, constitute important topics in scientometric research. In this research, collaboration is defined on the basis of (institutional) co-authorship. If all the authors of a paper are from one and the same institution, there is no formal institutional collaboration. If the authors are from two or more institutions located in the same country (and if no author from a foreign institution is involved), the collaboration is said to be national, while if the authors are from institutions located in two or more countries, the collaboration is defined as international.

**Torres Salinas** found for the *University of Navarra* that 26 per cent of its publication output resulted from international collaboration, 20 per cent from national collaboration, and 56 per cent did not result from collaboration formalized in institutional co-authorship. He found that internationally co-authored papers had on average a higher citation impact than papers resulting from national collabo-

ration, and that in turn nationally co-authored papers had a higher average impact than that of papers involving no collaboration. This outcome is in agreement with those obtained in similar studies for other universities.

Interestingly, however, **Torres-Salinas** further expanded this analysis by determining for each paper emerging from the *University of Navarra* the position of the Navarrese author(s) in the author list. Assuming that the first and the last author tend to be the most important contributors to a paper (i.e., the first author is often the junior researcher who carried out most of the work, whereas the last author is the project supervisor), he categorized the collaborative papers into two subclasses: those with a first or last author from the *University of Navarra* and those for which the author(s) from this university occupied an intermediate position (e.g., second or third author of a paper written by four authors). These two subclasses contained about 60 and 40 per cent of the collaborative papers published from the *University of Navarra*, respectively. Citation ratios were found both for internationally and for nationally co-authored papers.

He found that the average citation impact of an internationally co-authored paper in which the *University of Navarra* contributes the first or last author is lower than that of papers in which authors from this university occupied an intermediate position in the author list. This outcome is consistent with recent studies that are based on the notion

that in a statistical analysis of the effects of scientific collaboration and its relationship to citation impact, one should examine 'who is collaborating with whom' and also take into account the type of contribution an author or institution made to a collective effort. The outcomes presented by **Torres Salinas** are a clear illustration of this.

### 3. Response surface analysis

Scientific activity can be described in terms of an input-output model as a system with easily defined borders that transforms inputs, such as funding, research capacity and equipment, into outputs, such as publications or patents. The relationships that link inputs with outputs are complex. Therefore, tools must be applied that are capable of more complex modelling. **Torres Salinas** applied such a technique, called Response Surface Methodology. It emerged in the 1950s in chemical engineering in an attempt to construct empirical models able to find useful statistical relationships between all the variables making up an industrial system. In recent years it is being applied successfully in biology, medicine, and economics.

Using this methodology for a set of 22 research departments, the author analyzed the statistical relationship between a department's amount of human resources and funding on the one hand and the number of published articles on the other. In addition, he studied the relationship between the degree of collaboration and prestige of jour-

nals in which a department published its papers on the one hand, and the average citation impact of its papers on the other.

One may ask whether social processes as complex as the production of scientific knowledge or the scientific reward system are ruled by the same type of causality as chemical-technological processes. But the outcomes of the methodology are certainly of interest, and the conclusions significant. One of his findings is that departments that have a strong capacity to actively collaborate are capable of taking better advantage of the results of the research and tend to generate higher citation impacts. A second conclusion states that systems that have fewer human resources with better funding tend to be more productive than those with more human resources but less funding.

**Daniel Torres-Salinas'** dissertation has made substantial contributions not only to the design and technical realization of scientific information systems and to insight into the research performance of the *University of Navarra*, but also to the exploration of new data analysis techniques that potentially have a wider applicability and to a deeper understanding of key issues in the field of science and technology studies.

**Henk F. Moed**, Centre for Science and Technology Studies (CWTS), Leiden University, P. O. Box 9555, 2300 RB Leiden, the Netherlands. [moed@cwts.leidenuniv.nl](mailto:moed@cwts.leidenuniv.nl)

## PROMOCIÓN PARA NUEVOS SUSCRIPTORES DE EPI

Si te suscribes a **El profesional de la información** entre ahora y Navidad de 2007 te regalamos un *Anuario ThinkEPI 2007*.

Menciona esta oferta cuando te suscribas a través del formulario de suscripción online o del boletín en papel.